# Time-to-Event Analysis of Heart Failure via Electronic Health Records

Presenters:

Austin Falco

Aidan Kearns

Miranda Riggs

Tahmeed Tureen

STATS 404, 12/5/2017

# Agenda

- Background Information
- Purpose
- Conclusions
- Methods
- Assumptions
- Data
- Limitations
- Cohort Definition
- Results
- Significant Results
- Future Steps
- Acknowledgements
- Questions/Comments?

# Background Information

- Heart failure:
  - Occurs when the heart cannot pump blood to support the body's organs[1]
  - Doesn't mean the heart has stopped beating[1]

- According to Centers for Disease Control and Prevention (CDC)[1]:
  - Half of people diagnosed with heart failure die within 5 years
  - One in nine deaths in 2009 were at least partially attributed to heart failure

[1] https://www.cdc.gov/dhdsp/data_statistics/fact_sheets/fs_heart_failure.htm

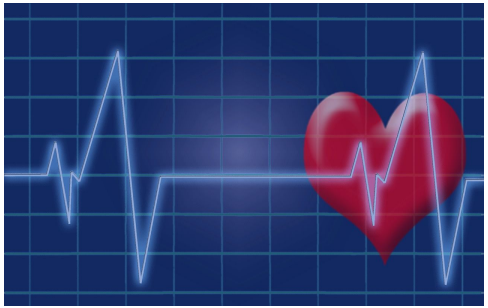
# Purpose

- Estimate and compare time until a heart failure diagnosis occurs
- Evaluate how certain risk factors affect time until heart failure diagnosis

https://pixabay.com/p-838655/?no_redirect

# Conclusions

- Significant risk factors differed between males and females
  - Males:
    - Cardiomyopathy
    - Severe Chronic Kidney Disease
  - Females:
    - Type 2 Diabetes
    - Sleep Disorder
- Some factors were common in both genders
  - Age
  - Hypertension



https://pixabay.com/p-1892826/?no_redirect

# Conclusions

The conclusions drawn from the project are consistent with readings from:

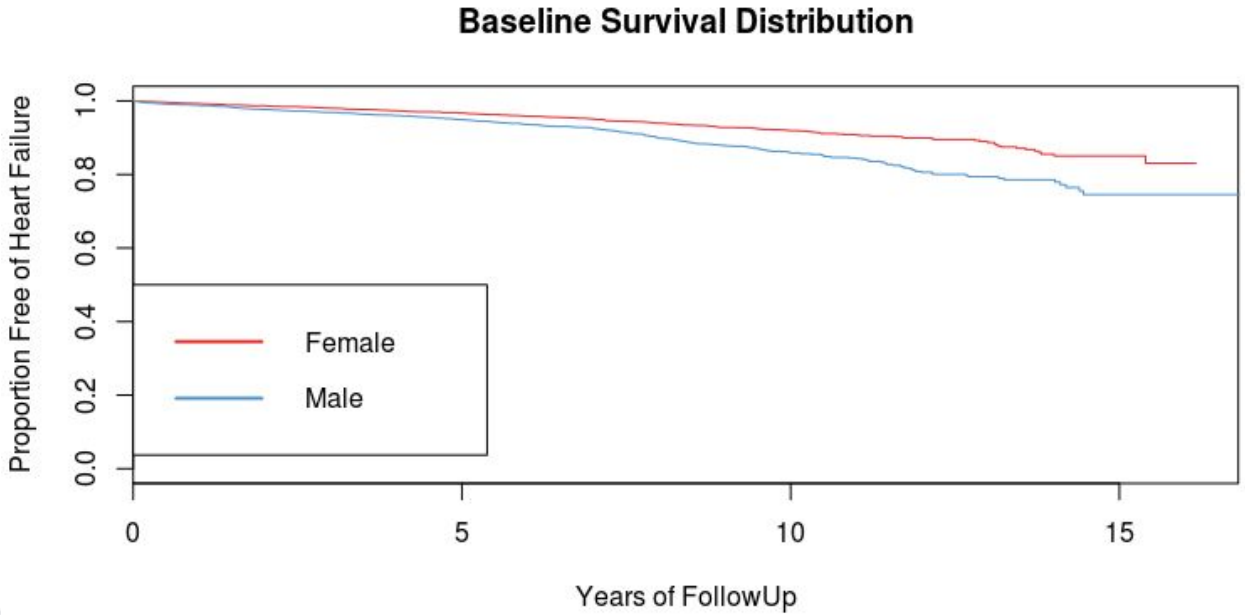
- National Institute of Health (NIH)
- Centers for Disease Control and Prevention (CDC)

# Questions? Comments?

# Methods

- Survival Analysis
  - Analyze the expected duration before an event occurs
  - Find proportion of the population that will survive past a given time
  - Observe characteristics that increase or decrease probability of survival
- Cox Proportional Hazard Model
  - Type of survival analysis
  - Works well for quantitative predictor variables (ex: Age)
  - Uses additional covariates (ex: Age, Arrhythmia, Cardiomyopathy, Type 2 Diabetes)
  - Estimates a hazards ratio

# Cox Proportional Hazard Model Example



# Assumptions

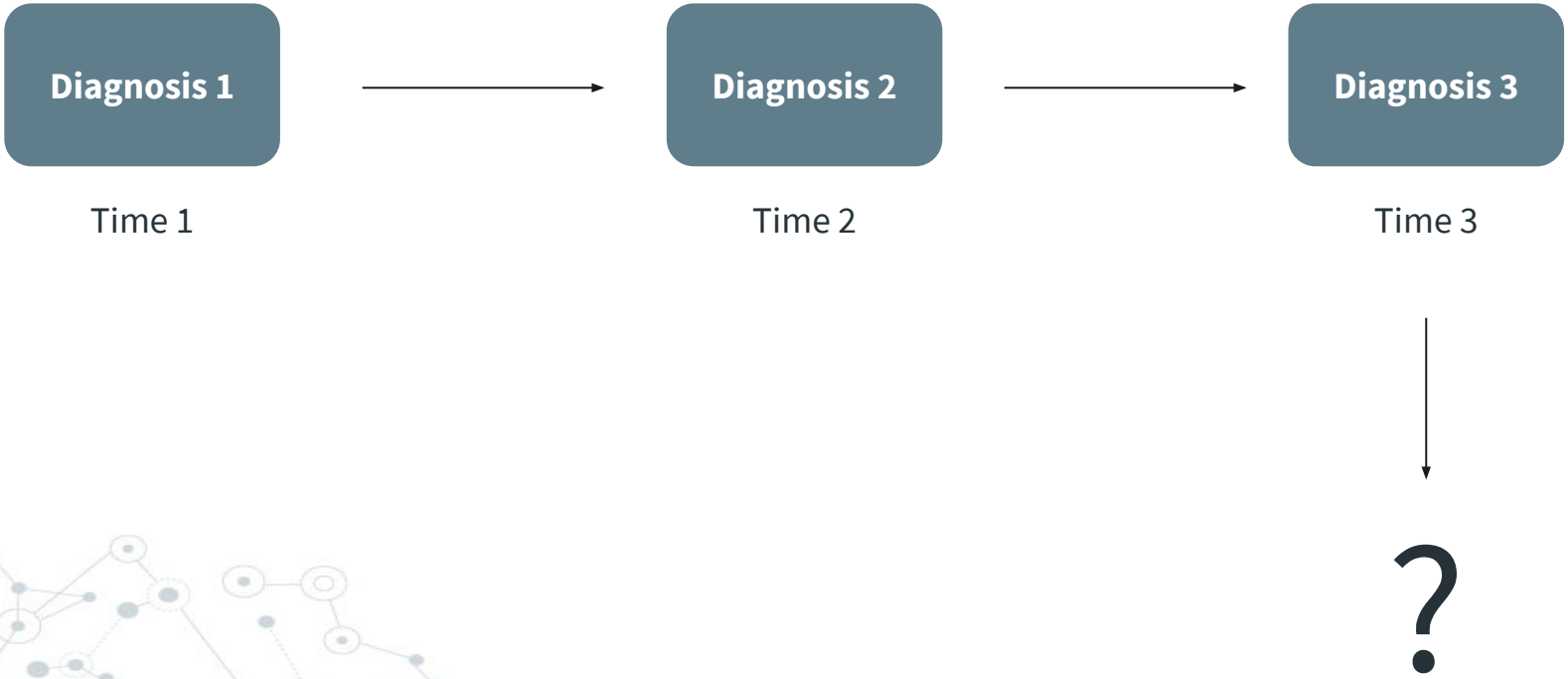
- Risk factors for heart failure chosen based on readings from Center for Disease Control and Prevention (CDC) and American Heart Association (AHA)
- University of Michigan health data representative of United States general population
- All diagnosis given by the University of Michigan Health System are accurate
- Cox Proportional Hazard predictor variables' effects on survival are constant over time

# Questions? Comments?

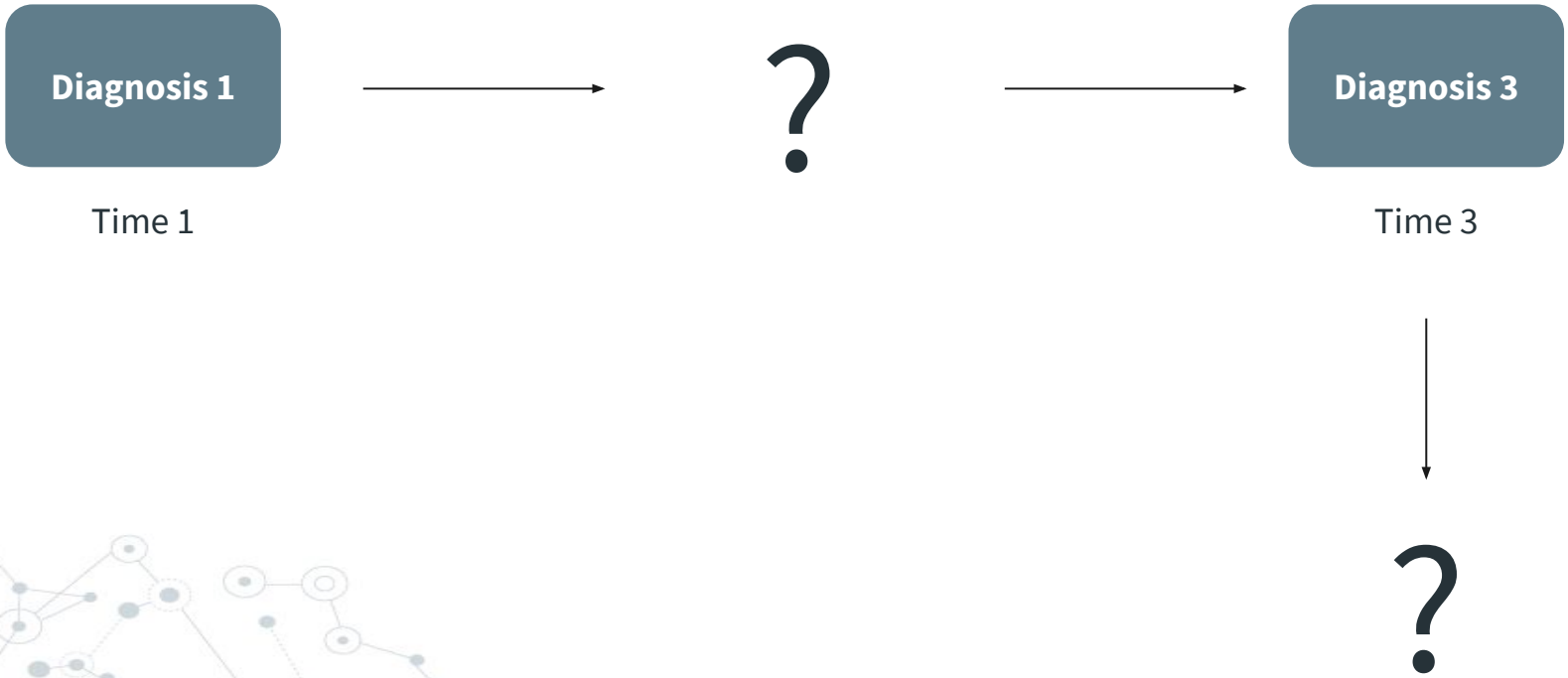
# Data - Electronic Health Records (EHR)

- Electronically-stored health information of patients
- A single EHR is a diagnosis of a patient on a given date that is stored in a hospital/medical system
- For this study, the EHR dataset was provided by the University of Michigan Health System

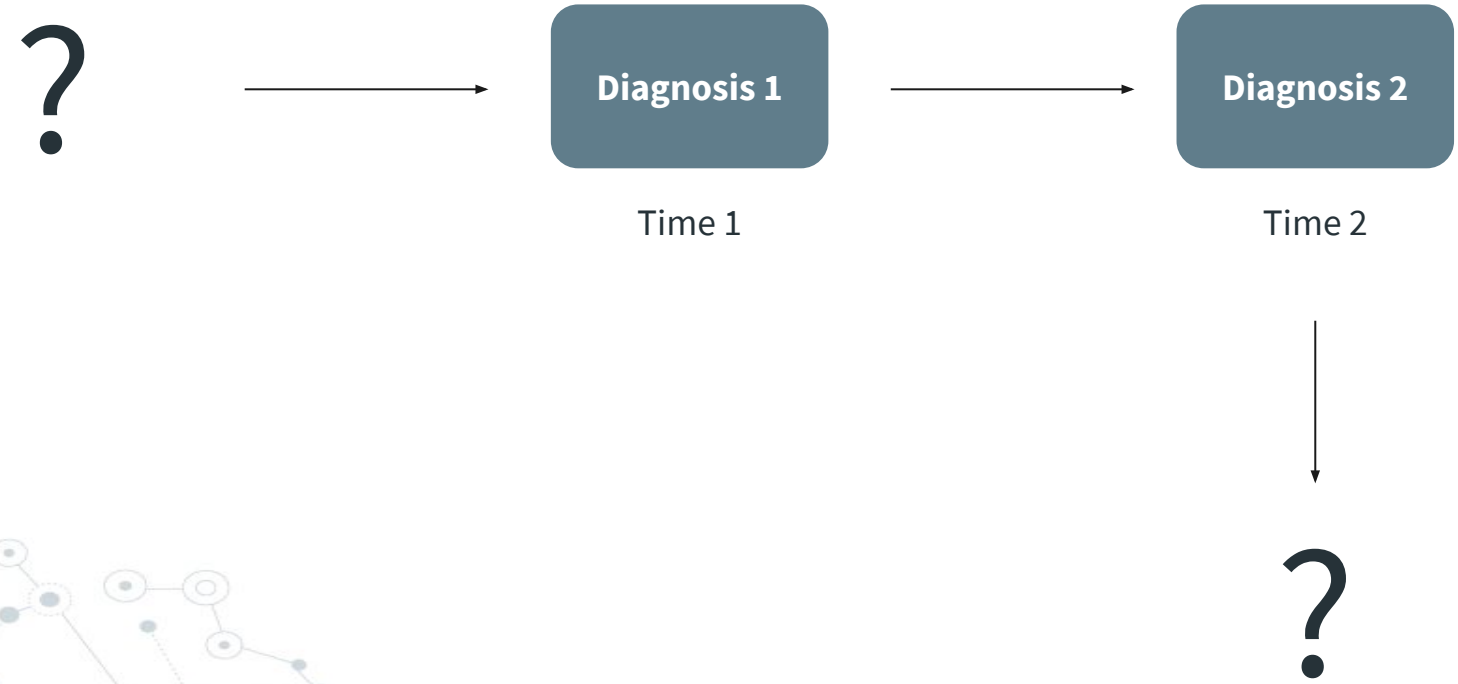
# How EHR behaves ...1



# How EHR behaves ...2



# How EHR behaves ...3



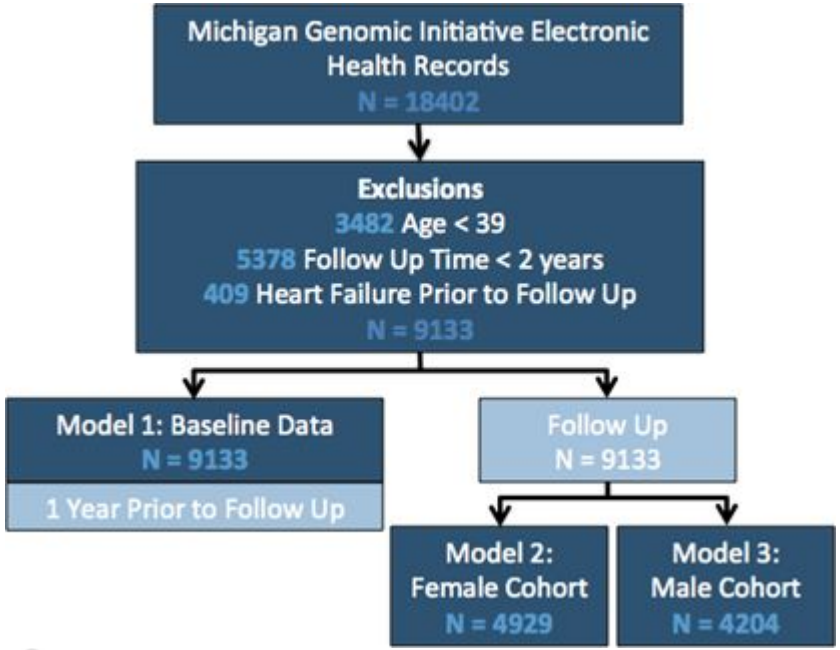
# Important because

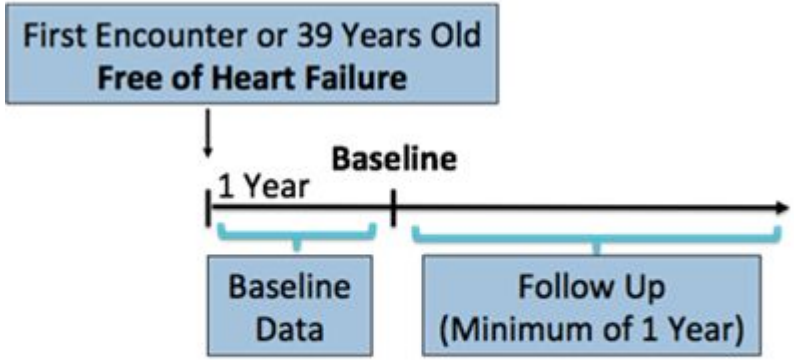
- EHR
  - Incomplete
  - Non-continuous
- Influence
  - Estimates from survival analysis

# Limitations

- Textual data differences in diagnosis
- Disease diagnosis codes changed after 2014
- Large quantities of missing data

# Cohort Definition (Analysis dataset)




Outcome: Heart Failure as defined by ICD 9 (428) and ICD 10 (I50)

# Questions? Comments?

# Results

**TABLE 1.** Hazard ratios for male cohort

| Risk Factor | Hazard Ratio |
|---|---|
| Age | 1.052* |
| Cardiomyopathy | 10.203* |
| Hypertension | 1.337* |
| Severe Chronic Kidney Disease | 3.08* |
| Type 2 Diabetes Mellitus | 3.526 |
| Emphysema | 1.684 |
| Moderate Chronic Kidney Disease | 1.725 |
| Hepatitis C | 1.945 |
| Arrhythmia | 1.259 |
| Type 2 Diabetes Mellitus*Age | 0.987 |
| Mild Chronic Kidney Disease | 0.569 |
| Hypercholesterol | 0.911 |
| Failure Chronic Kidney Disease | 1.191 |
| Sleep Disorder | 1.036 |
| Morbid Obesity | 1.064 |

Source: Michigan Genomics Initiative, University of Michigan

*Note:* n = 4204

*p < .05

# Results

**TABLE 2.** Hazard ratios for female cohort

| Risk Factor | Hazard Ratio |
|---|---|
| Age | 1.052* |
| Type 2 Diabetes Mellitus | 9.350* |
| Hypertension | 1.442* |
| Sleep Disorder | 1.764* |
| Failure Chronic Kidney Disease | 4.179 |
| Type 2 Diabetes Mellitus*Age | 0.973 |
| Arrhythmia | 1.563 |
| Morbid Obesity | 1.635 |
| Hepatitis C | 3.474 |
| Moderate Chronic Kidney Disease | 2.169 |
| Severe Chronic Kidney Disease | 2.070 |
| Hypercholesterol | 0.776 |
| Mild Chronic Kidney Disease | 1.675 |
| Emphysema | 1.333 |
| Cardiomyopathy | 0.967 |

Source: Michigan Genomics Initiative, University of Michigan

*Note:* n = 4929

*p < .05

# Significant Results

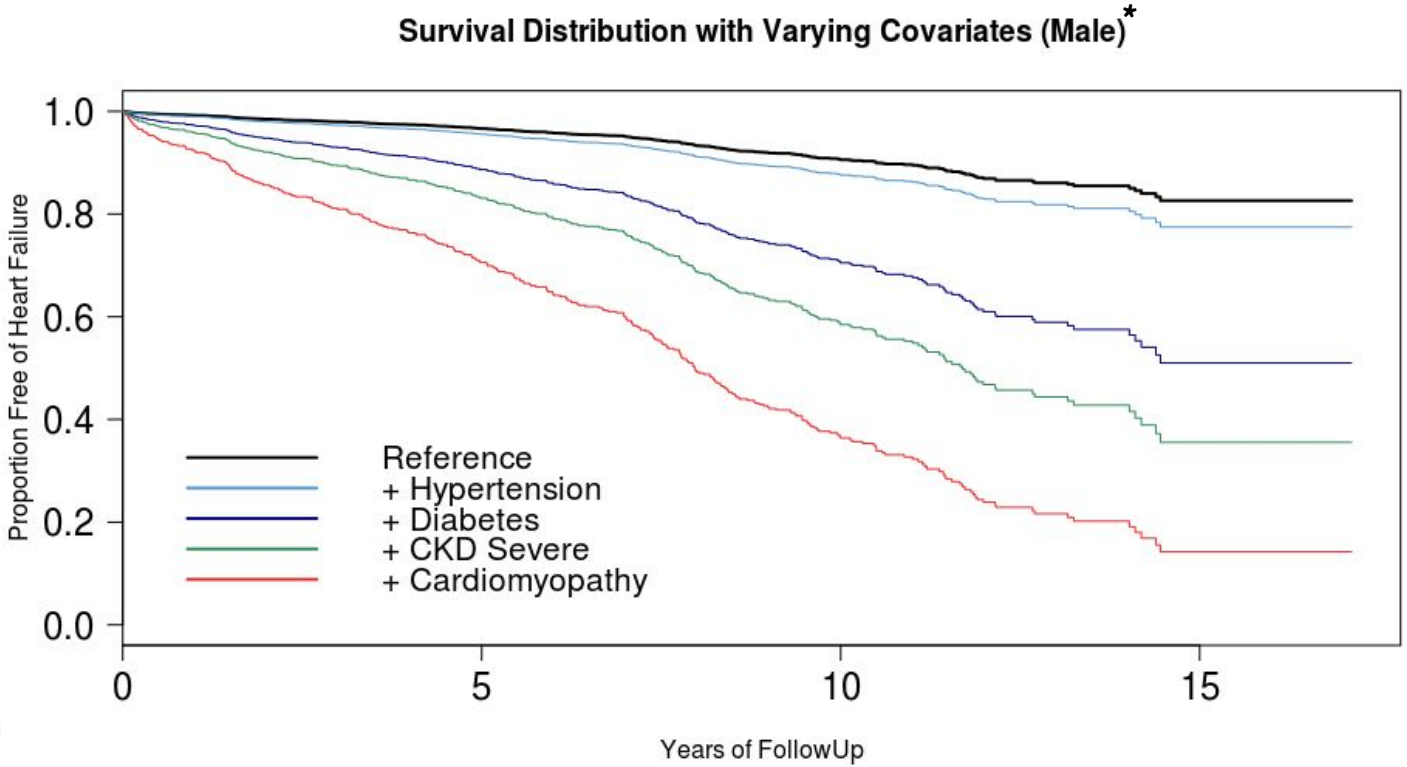
**TABLE 3.** Significant hazard ratios for male cohort

| Risk Factor | Hazard Ratio |
|---|---|
| Age | 1.052 |
| Cardiomyopathy | 10.203 |
| Hypertension | 1.337 |
| Severe Chronic Kidney Disease | 3.080 |

Source: Michigan Genomics Initiative, University of Michigan

*Note:* n = 4204

# Significant Results




** Baseline patient is 60 years old and has been diagnosed with Arrhythmia, Mild CKD, and Hypercholesterol*

# Significant Results

**TABLE 4.** Significant hazard ratios for female cohort

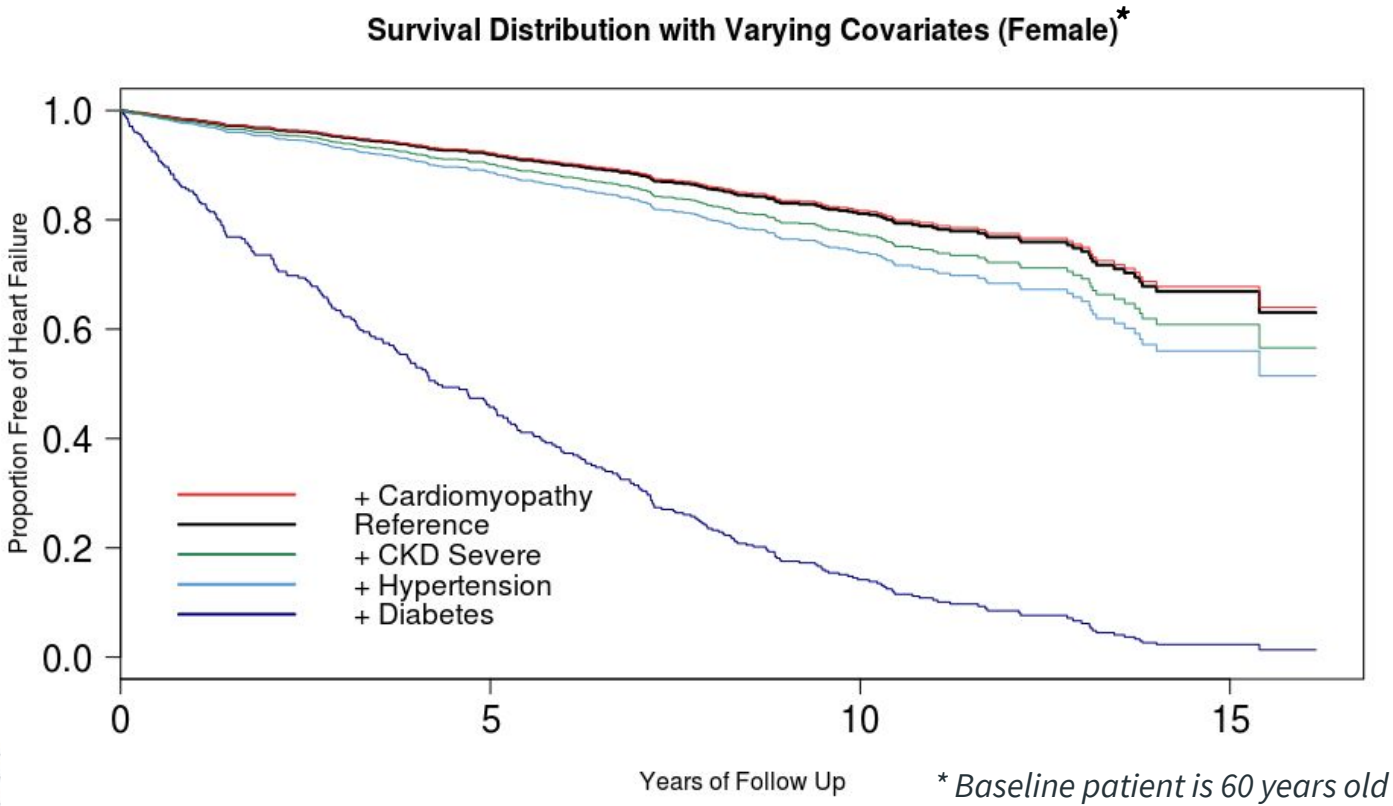
| Risk Factor | Hazard Ratio |
|---|---|
| Age | 1.052 |
| Type 2 Diabetes Mellitus | 9.350 |
| Hypertension | 1.442 |
| Sleep Disorder | 1.764 |

Source: Michigan Genomics Initiative, University of Michigan

*Note:* n = 4929

*Note:* Type 2 Diabetes Mellitus had a 95% confidence interval of (1.433, 61.024)

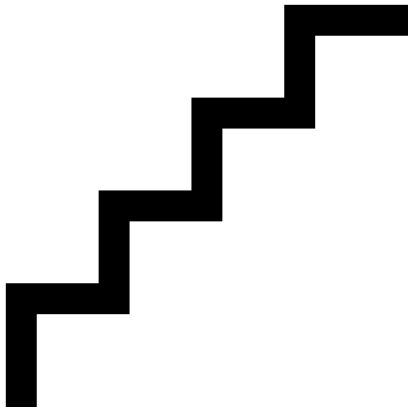
# Significant Results




** Baseline patient is 60 years old and has been diagnosed with Arrhythmia, Mild CKD, and Hypercholesterol*

# Future Steps

- Explore shrinkage methods
- Analyze significant risk factors
- Discuss research potential of EHR



https://upload.wikimedia.org/wikipedia/commons/thumb/2/2c/Steps.svg/2000px-Steps.svg.png

# Research Acknowledgements


**Deesha Bhaumik**

*Undergraduate Student, Public Health, Johns Hopkins University*

**Christian Erickson**

*Undergraduate Student, Mathematics, Hope College*

**Tahmeed Tureen**

*Undergraduate Student, Statistics, University of Michigan*

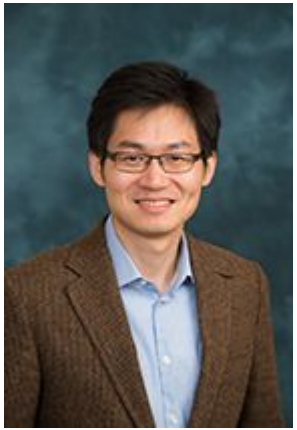
# Faculty Acknowledgements

**Dr. Philip Boonstra**

*Assistant Professor*

**Dr. Zhenke Wu**

*Assistant Professor*



**Dr. Matthew Zawistowski**

*Research Specialist*

Department of Biostatistics, University of Michigan

# Questions? Comments?

# Thank you!


**Austin Falco**

Undergraduate Student
Department of Statistics
University of Michigan

aufalco@umich.edu

**Aidan Kearns**

Undergraduate Student
Department of Statistics
University of Michigan

aidank@umich.edu

**Miranda Riggs**

Undergraduate Student
Department of Statistics
University of Michigan

riggsmir@umich.edu

**Tahmeed Tureen**

Undergraduate Student
Department of Statistics
University of Michigan

tureen@umich.edu

# Big Data Summer Institute, University of Michigan

- https://sph.umich.edu/bdsi/
- Supervised by Dr. Bhramar Mukherjee, Dept. of Biostatistics
- Students interested in statistical research in health
- Students interested in graduate studies in mathematics, statistics, computer science and biostatistics

For more info,

Contact Tahmeed Tureen